IS ARTIFICIAL INTELLIGENCE POSSIBLE?

Project Specification DD143X

Author: Rishie Sharma

Coordinator: Johan Boye

2011-02-08

SUMMARY

In this document some expectations and possibilities of artificial intelligence are briefly discussed. While the field of artificial intelligence has progressed substantially since its birth it still has a long way to fulfil the dreams of science fiction. The question is, is it even possible to fulfil those dreams, and if so, how will it be made possible? This project's aim is to try and answer these questions in the best possible way.

BACKGROUND

Ever since its inception, artificial intelligence has had a wondrous connotation to it. People think of robots taking over the world, super-intelligent machines incapable of error, technological singularities and the like. Even back in the 1930s Alan Turing had grand dreams for the future and predicted:

"I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted."

Already 11 years have passed since the end of that century and no one in their right mind would claim that any current existing machine is capable of thinking like humans. Even if we have doors that open themselves when a person approaches, computers winning over grand masters in chess, navigation systems that tell a car driver what turn to make, aeroplanes more or less flying themselves and many more automated systems, the fact remains that none of them have anything that resembles human thinking. Artificial intelligence, in the sense of creating intelligence similar to our own, proved to be harder than expected (by Turing if by no other).

Some philosophers tend to argue about *strong* versus *weak* AI. The strong AI hypothesis ponders the question if it is possible to create an artificial consciousness, while the idea that machines can be made to mimic the intelligent behaviour of humans is called weak AI. Almost everyone agrees that weak AI is possible and any engineer would agree that it is the only relevant question to ask.

In any given application it has already been done or is quite easy to accept the thought that in the future, a computer program will be able to perform as well or even better than a human. This is the idea of an intelligent agent and the study of artificial intelligence is usually said to be the study of intelligent agents that perceive their environment and take actions that maximize their chances of success. There are many areas where humans greatly excel over computer programs such as playing Go, translating text between languages, describing pictures with words among others. However, it is not such a far-fetched idea that computer programs will improve with time as research in these areas advance to finally in the future surpass the human capability. Still it is questionable to say that these intelligent agents would really be intelligent, one would be inclined to say that they are just automated systems, possessing no "real" intelligence.

But then what is intelligence, and thus in extension, artificial intelligence? As a materialist one would want to say that our intelligence is merely an illusion created by the extreme complexity of our brain. The combined powers of our instincts, emotions, experiences and well evolved heuristics of our brain are what constitute our intelligence. Phrases such as "to err is human" and "perfection comes through repetition" gives us an idea of how the human

mind works, and it is nothing like the classical logic based computer programs. However, with the advent of artificial neural networks and projects like The Blue Brain Project we are beginning to closer simulate the processing power of the brain and perhaps closer to simulating “real” intelligence.

Disregarding the ideas of strong or weak AI let us return to the notion of creating something resembling human intelligence. Turing proposed to replace the question “Can machines think?” with a test in which the principle that if machine intelligence is indistinguishable from human intelligence, then the former is as “good” as the latter. While a machine capable of consistently fooling a human that it is a human would be a truly remarkable feat, there could also be machines not capable of this but still worthy of the title artificial intelligence.

STATEMENT OF THE PROBLEM

I consider the question if computers can perform as well as humans in specific tasks trivial except perhaps in extremely subjective tasks such as suggesting improvements to the composition of a piece of music. Moreover, those kinds of tasks have no correct answers anyway. Consequently I will investigate the possibility of one coherent machine system matching or even outperforming most humans at most intellectual tasks. With a coherent machine system I mean one where no human interaction is needed except speaking to it. You would not need to enter the system and run the chess bot when you want to play a game of chess but rather all details would be hidden to you and you would only interact with it through natural language.

PROJECT PLAN

A couple of definitions need to be made. Define more precisely what a coherent machine system would be, perhaps redefining the statement of the problem accordingly when a more accurate, well thought-out description is found. Secondly, define what matching or even outperforming most humans means more precisely and lastly define intellectual tasks.

With everything well-defined, discuss if such a system is possible and if so how such a system could be created and what problems could occur when doing so.

TIME SCHEDULE

- Project specification — 9 February
- **Have precise definitions all ready** — **1 March**
- Halfway meeting — 4 March
- Essay hand in deadline — 14 April
- Review hand in deadline — 26 April
- Exjobb conference — 27 April
- Final version of essay — TBA

REFERENCES

A. Turing: "Computing machinery and intelligence", http://www.abelard.org/turpap/turpap.php (originally published in "Mind" 1950). The seminal paper in artificial intelligence; very readable. Contains the first description of the "Turing test" (called the "imitation game" in the article).

J. Searle: "Minds, brains and programs". http://web.archive.org/web/20071210043312/http://members.aol.com/NeoNoetics/MindsBrainsPrograms.html (originally published in Behavioral and Brain Sciences). Contains the original Chinese room argument.

S. Russell and P. Norvig. "Artificial intelligence -- A modern approach", chapters 1 and 26. Prentice-Hall, 1995.